We claim:

1. A method of synthesizing a repertoire of oligonucleotide tags of a predetermined length, the method comprising the steps of:

    (a) providing a repertoire of oligonucleotide tag precursors in an amplicon, the oligonucleotide tag precursors each comprising one or more words, and each of the one or more words being selected from the same minimally cross-hybridizing set;

    (b) cleaving the amplicon at a word in each of the oligonucleotide tag precursors to form one or more ligatable ends on each oligonucleotide tag precursor;

    (c) ligating one or more words to the one or more ligatable ends to elongate each of the oligonucleotide tag precursors;

    (d) amplifying the elongated oligonucleotide tag precursors in the amplicon; and

    (e) repeating steps (b) through (d) until a repertoire of oligonucleotide tags having the predetermined length is formed.

2. The method of claim 1 wherein said amplicon is a cloning vector.

3 The method of claim 2 wherein said step of cleaving includes cleaving said amplicon in a region adjacent to said word by a type IIs restriction endonuclease.

4. The method of claim 3 wherein said word has a length in the range of from three to fourteen nucleotides.

5. The method of claim 4 wherein oligonucleotide tag has a length in the range of from 18 to 60 nucleotides.

6. The method of claim 2 wherein said step of cleaving includes cleaving said amplicon across said word by a type IIs restriction endonuclease.

7. The method of claim 2 wherein said word has a length of four and wherein said oligonucleotide tag has a length in the range of from 18 to 40.

8. A repertoire of oligonucleotide tags, wherein the oligonucleotide tags of the repertoire are of the form:

$$w_1(N)_{x1}w_2(N)_{x2} \ldots (N)_{xn-1}w_n$$

- 31 -

wherein each of w_1 through w_n is a word consisting of an oligonucleotide having a length from three to fourteen nucleotides or basepairs and being selected from the same minimally cross-hybridizing set wherein a word of the set and a complement of any other word of the set has at least two mismatches; N is a nucleotide or basepair; each of x_1 through x_{n-1} is an integer selected from the group consisting of 0, 1, 2, 3, and 4, provided that at least one of x_1 through x_{n-1} is 1, 2, 3, or 4; and n is an integer in the range of from 4 to 10.

9. The repertoire of claim 8 wherein each of said x_1 through x_{n-1} is selected from the group consisting of 0, 1, and 2, and wherein said length of said word is from four to ten nucleotides or basepairs.

10. The repertoire of claim 9 wherein said oligonucleotide tags are single stranded and wherein n is in the range of from 6 to 10.

11. The repertoire of claim 10 wherein a duplex between each of said words of said minimally cross-hybridizing set and said complement of any other word of said set would have at least three mismatches.

12. The repertoire of claim 11 wherein a duplex between each of said words of said minimally cross-hybridizing set and said complement of any other word of said set would have at least five mismatches whenever said word has a length of greater than or equal to six nucleotides.

13 The repertoire of claim 10 having a number of said oligonucleotide tags that is in the range of from 100 to 1×10^9.

14. The repertoire of claim 13 having a number of said oligonucleotide tags that is in the range of from 1000 to 1×10^8.

15. A repertoire of cloning vectors for attaching oligonucleotide tags to polynucleotides, wherein each of the vectors comprises a double stranded element corresponding to an oligonucleotide tag of the form:

$$w_1(N)_{x1}w_2(N)_{x2} \cdots (N)_{xn-1}w_n$$

wherein each of w_1 through w_n is a word consisting of an oligonucleotide having a length from three to fourteen nucleotides and being selected from the same minimally cross-hybridizing set wherein a word of the set and a complement of any other word of the set has at least two

mismatches; N is a nucleotide; each of x_1 through x_{n-1} is an integer selected from the group consisting of 0, 1, 2, 3, and 4, provided that at least one of x_1 through x_{n-1} is 1, 2, 3, or 4; and n is an integer in the range of from 4 to 10.

5   16.   The repertoire of claim 15 wherein each of said x_1 through x_{n-1} is selected from the group consisting of 0, 1, and 2, and wherein said length of said word is from four to ten nucleotides or basepairs.